CLAIMS

1   1.   A computer system, comprising:

2   a general purpose computer, the computer including logic for undertaking method acts

3   to map data arranged in a source schema into a target schema, the method acts undertaken

4   by the logic including:

5   receiving at least one value correspondence, each value correspondence representing

6   a function for deriving a value of a target attribute from one or more values of source

7   attributes;

8   grouping at least some value correspondences into potential sets;

9   selecting candidate sets from at least some potential sets;

10   grouping at least some candidate sets into covers; and

11   using at least one cover, generating at least one query representing a source schema-to-

12   target schema mapping.

1   2.   The computer of Claim 1, wherein the method acts undertaken by the logic to execute

2   the grouping act include:

3   grouping value correspondences into potential sets such that, for each potential set, at

4   most one value correspondence per target attribute exists.

1   3.   The computer of Claim 2, wherein the method acts undertaken by the logic further

2   include:

3   adding a potential set to a set of candidate sets if only one source relation is used to

4   compute mappings using the potential set; otherwise

5   adding a potential set to the set of candidate sets only if a join path for the source

6   relations can be identified.

1   4.   The computer of Claim 3, wherein the method acts undertaken by the logic further

2   include:

3   arranging candidate sets into groups such that each group includes every value

4   correspondence at least once, the groups establishing covers.

1   5.   The computer of Claim 4, wherein the method acts undertaken by the logic further

2   include:

3   establishing at least one selected cover;

          for each candidate set in the selected cover, creating at least one query; and

4

5   combining the queries for the cover.

1   6.   The computer of Claim 5, wherein the method acts undertaken by the logic to establish

2   at least one selected cover include:

3   ranking the covers by at least one of: a number of candidate sets in each cover, and

4   a number of target attributes; and

5   presenting ranked covers to a user for selection of one of the covers as the selected

6   cover.

1   7.   The computer of Claim 3, wherein the logic undertakes the act of adding a potential

2   set to the set of candidate sets only if a join path for the source relations can be identified using a

3   spanning tree.

1   8.   The computer of Claim 1, wherein the logic incrementally undertakes the acts of

2   grouping value correspondences into potential sets, selecting candidate sets, grouping candidate sets

3   into covers, and generating queries representing mappings.

1   9.   A computer-implemented method for generating a mapping from a source schema to

2   a target schema, comprising:

3   generating a mapping based on at least a subset of value correspondences, each value

4   correspondence representing a function for deriving a value of a target attribute from one or

5   more values of source attributes;

6   allowing a user, in a user interaction, to incrementally add or delete a value

7   correspondence from the subset;

8   based on the user interaction, generating a new mapping;

9   presenting a representation of the new mapping to the user such that the user can view

10   the representation; and

11   permitting the user to add or delete a value correspondence embodied in the new

12   mapping to generate another mapping.

1 10. The method of Claim 9, wherein the generating act includes:

2 grouping at least some value correspondences into potential sets;

3 selecting candidate sets from at least some potential sets;

4 grouping at least some candidate sets into covers; and

5 using at least one cover, generating at least one query representing a source schema-to-

6 target schema mapping.

11. The method of Claim 10, further comprising:

grouping value correspondences into potential sets such that, for each potential set, at

most one value correspondence per target attribute exists.

12. The method of Claim 11, further comprising:

adding a potential set to a set of candidate sets if only one source relation is used to

compute mappings using the potential set; otherwise

4 adding a potential set to the set of candidate sets only if a join path for the source

5 relations can be identified.

1 13. The method of Claim 12, further comprising:

2 arranging candidate sets into groups such that each group includes every value

3 correspondence at least once, the groups establishing covers.

1 14. The method of Claim 13, further comprising:

2   establishing at least one selected cover;

3   for each candidate set in the selected cover, creating at least one query; and

4   combining the queries for the cover.

1   15. The method of Claim 14, further comprising:

2   ranking the covers by at least one of: a number of candidate sets in each cover, and

3 a number of target attributes; and

4   presenting ranked covers to a user for selection of one of the covers as the selected

5 cover.

1   16. The method of Claim 12, further comprising adding a potential set to the set of

2 candidate sets only if a join path for the source relations can be identified using a spanning tree.

1   17. A computer program device comprising:

2   a computer program storage device readable by a digital processing apparatus; and

3   a program on the program storage device and including instructions executable by the digital

4 processing apparatus for performing method acts for generating a query representing a source

5 schema-to-target schema mapping, the program comprising:

6   computer readable code means for establishing plural value correspondences;

7   computer readable code means for generating subsets of value correspondences such

8 that (1) each subset has at most one value correspondence per target attribute, (2) for each

9 subset requiring more than one source relation to undertake a mapping, a join path can be

10   found between the relations, and (3) each subset includes at least every value correspondence;

11   and

12   computer readable code means for generating a query using one of the subsets, the

13   query being representative of a source schema-to-target schema mapping.

1   18.   The computer program product of Claim 17, further comprising computer readable

2   code means for sorting the subsets and displaying at least portions of a sorted list of subsets, such

3   that a user can establish a selected subset used to generate the query.

1   19.   The computer program product of Claim 18, wherein the means for generating subsets

2   generates candidate sets, each subset including one or more candidate sets, and the means for sorting

3   sorts the subsets by inverse number of candidate sets.

1   20.   The computer program product of Claim 19, wherein the means for sorting also sorts

2   the subsets by the number of value correspondences in the subsets.

1   21.   The computer program product of Claim 19, wherein the means for generating a query

2   creates at least one query for each candidate set in the selected subset, and then combines the queries

3   for the subset.

1   22.   The computer program product of Claim 21, wherein the means for generating subsets

2   and the means for generating a query are incrementally invoked by a user to selectively add or delete

3   value correspondences from a selected subset.